SEQUENCE LISTING

<110> Gellissen, Gerd
 Braus, Gerhard
 Pries, Ralph
 Krappmann, Sven
 Strasser, Alexander

<120> Nucleic Acid Molecule Comprising a Nucleic Acid Coding for a Polypeptide with Chorismate Mutase Activity

<130> 029474-5007-00 <140> 10/042059 <141> 2001-10-25 <150> DE 199 19 124.7 <151> 1999-04-27 <160> 7 <170> PatentIn version 3.1 <210> <211> 843 <212> DNA <213> Hansenula polymorpha <220> <221> CDS <222> (1)..(843) <223> <400> 1 atg gac ttt atg aag cca gaa aca gtg ctg gac ctt ggc aac att aga 48 Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg 10 gat gcc ttg gtc cgg atg gag gat acg atc atc ttc aac ttt atc gag 96 Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu 20 cgg tcg cag ttc tat gcg tcg ccc tcg gta tac aaa gtc aac cag ttc 144 Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe 35 cct att ccc aac ttc gac ggc tcg ttc ttg gac tgg ctg ttg tcg cag 192 Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln 50 55 cac gag cga atc cat tcg cag gtg agg aga tac gac gcg cca gac gag 240 His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu 288 gtg cct ttt ttc ccc aac gtg ctg gaa aaa acg ttt ctg ccc aag atc Val Pro Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile 85 90



						gcc Ala										336
						tac Tyr										384
_		_			_	gag Glu 135	_				_	_	_	_	_	432
_			-	_	_	tcg Ser			_	_						480
	_	_		_		ttt Phe		-			_	_			_	528
_			_	_	_	gtg Val	_									576
_		_	_			atc Ile	_	_	_		_		_			624
				_	_	aca Thr 215		_		_	_			-	_	672
						att Ile										720
_		_	_	_	_	gaa Glu	_	_		_	_	_		_		768
						gcg Ala										816
						ttg Leu		tag								843
		_														

<210> 2

<211> 280 <212> PRT <213> Hansenula polymorpha

<400> 2

Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg 5 Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln 55 His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu Val Pro Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile 85 90 Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn 100 105 Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala 115 120 Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala 130 135 Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg 150 155 Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp 165 170 Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn 180 185 Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg 195 200 205 Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser 210 215 220 Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile

225	230	235	240

Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu 245 250 255

Asp Glu Glu Asp Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp 260 265 270

Arg Phe Leu Ser Ser Gly Leu Tyr 275 280

<210> 3

<211> 1655

<212> DNA

<213> Hansenula polymorpha

<220>

<221> gene

<222> (1)..(1655)

<223> 1,8 kb genomic DNA-fragment from Hanseula polymorpha

<220>

<221> gene

<222> (1)..(1655)

<223> 1,8 kb genomic DNA-fragment from Hansenula polymorpha

<400> 3

eccggeecaa tgecageaat atggagaegt ttaggeagaa taggegttee ataettetea 60 cgctgcttgt tgccaccgga atatacaccg cattgcagtt tgcacacatc atactatatg 120 acgattacat tggcggaacg tatcgcgagt cgctcacgag acgcattaga atgacagaga 180 aatcgcgaaa cgaccttata gacgcacgtg aaaactacgg gtttggaggc agcaaggagg 240 agcgaatcca gcggtttttg tggttcagac atctttcgtg gcttttaggc gaggataagc 300 gaacttgagg agcgtttttt ttttcctgtt tagtttttgt aggtatggac tttatgaagc 360 cagaaacagt gctggacctt ggcaacatta gagatgcctt ggtccggatg gaggatacga 420 tcatcttcaa ctttatcgag cggtcgcagt tctatgcgtc gccctcggta tacaaagtca 480 accagttccc tattcccaac ttcgacggct cgttcttgga ctggctgttg tcgcagcacg 540 agegaateea ttegeaggtg aggagataeg aegegeeaga egaggtgeet ttttteeeea 600 acgtgctgga aaaaacgttt ctgcccaaga tcaactaccc atcggtgcta gcctcctacg 660 cggatgaaat caacgtcaac aaagagatac tcaagatcta cacgtcagag atagtaccag 720

gaatagctgc	aggcagcgga	gagcaggagg	acaaccttgg	ctcgtgcgca	atggccgaca	. 780
tcgagtgcct	gcagtcgcta	tccagaagaa	tccattttgg	ccgttttgtc	gcagaggcta	840
aatttatcag	tgagggggac	aagattgtgg	atctgatcaa	aaagagagat	gtggaaggca	900
ttgaggcgct	catcacaaac	gccgaggtcg	aaaaacggat	cttggacaga	cttctggaga	960
agggaagggc	gtatggaaca	gacccgacac	taaagttcac	gcagcacatt	cagagcaagg	1020
tgaagcccga	ggtgattgtg	aaaatctaca	aggatttcgt	gattccgctc	acgaagaagg	1080
tcgaagtcga	ctacttgctg	agacggctgg	aggacgagga	ggacgatgat	gcgacgcaga	1140
aaagcggcgg	ctacgttgac	cggtttctct	cctctggctt	gtactagaaa	ttaaaatttt	1200
cagtacttta	attattctcg	aattctagtt	cagataccgc	atggtaattt	caaaggccag	1260
aaaagtggcc	gcgttggctg	gggcagctct	cagaatagtc	ggcgagaatc	ctttgactag	1320
cccccaggca	ccgctctgtc	tccaaatacc	cctaatagtc	tcaacagcat	ttctataaac	1380
cagcttcttg	tagttgtccg	tctgcatgtt	ggacttgatc	acatcgatcg	gataaatact	1440
gaaccacatc	ccgtaacctg	ccagcgcccc	aaagacgcag	agcttccagt	tctcgatgtc	1500
cttcctggca	atattccgcg	actcgatctc	gtttttcacg	agagcttcaa	aagtcagaaa	1560
atacgctccg	ctacccaaac	tttctcttgc	cagcgtaggt	cccagacccc	ggtagattaa	1620
cttgatgcct	cccgtatggt	acagcttctt	gatee			1655

<210> 4

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotide

<400> 4

aattaaccct cactaaaggg

20

<210> 5

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotide

<400> 5

gtaatacgac tcactatagg gc

22

<210>	6	
<211>	26	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	Oligonucleotide	
<400>		
atataga	atct acaaaaacta aacagg	26
010		
<210>		
<211>		
<212>		
<213>	Artificial	
<220>		
<223>	Oligonucleotide	
.400-		
<400>		
atataga	atct gatgcgacgc agaaaagc	28

U.S. Patent Application No. 10/042,059 of Gellissen et al. Attorney Docket No. 029474-5007-00 Marked-up Sequence Listing

SEQUENCE LISTING

<110> Gellissen, Gerd Braus, Gerhard Pries, Ralph Krappmann, Sven Strasser, Alexander

<120> Nucleic Acid Molecule Comprising a Nucleic Acid Coding for a Polypeptide with Chorismate Mutase Activity

with	Cn	orısı	nate	Muta	ise A	ACTIV	rty									
<130	<130> 029474-5007-00															
<140 <141	·															
<150 <151		DE 199 19 124.7 1999-04-27														
<160	> .	7														
<170	>	Pater	ntIn	vers	sion	3.1										
<211 <212	<210> 1 <211> 843 <212> DNA <213> Hansenula polymorpha															
<221 <222	<220> <221> CDS <222> (1)(843) <223>															
	gac	1 ttt Phe														48
gat Asp	gcc Ala	ttg Leu	gtc Val 20	cgg Arg	atg Met	gag Glu	gat Asp	acg Thr 25	atc Ile	atc Ile	ttc Phe	aac Asn	ttt Phe 30	atc Ile	gag Glu	96
		cag Gln 35														144
		ccc Pro														192
		cga Arg														240
		ttt Phe														288
		cca Pro														336

100	105	1
100	 105	-

			ata gta cca gga Ile Val Pro Gly 125						
gca ggc agc gg Ala Gly Ser Gl 130	ga gag cag ga Ly Glu Gln Gl 13	u Asp Asn Leu	ggc tcg tgc gca Gly Ser Cys Ala 140	atg gcc 432 Met Ala					
gac atc gag to Asp Ile Glu Cy 145	gc ctg cag to ys Leu Gln Se 150	g cta tcc aga r Leu Ser Arg	aga atc cat ttt Arg Ile His Phe 155	ggc cgt 480 Gly Arg 160					
			ggg gac aag att Gly Asp Lys Ile						
ctg atc aaa aa Leu Ile Lys Ly 18	ys Arg Asp Va	g gaa ggc att l Glu Gly Ile 185	gag gcg ctc atc Glu Ala Leu Ile 190	Thr Asn					
			ctt ctg gag aag Leu Leu Glu Lys 205						
gcg tat gga ac Ala Tyr Gly Th 210	ca gac ccg ac nr Asp Pro Th 21	r Leu Lys Phe	acg cag cac att Thr Gln His Ile 220	cag agc 672 Gln Ser					
aag gtg aag co Lys Val Lys Pr 225	cc gag gtg at ro Glu Val Il 230	t gtg aaa atc e Val Lys Ile	tac aag gat tto Tyr Lys Asp Pho 235	gtg att 720 Val Ile 240					
ccg ctc acg as Pro Leu Thr Ly	ag aag gtc ga ys Lys Val Gl 245	a gtc gac tac u Val Asp Tyr 250	ttg ctg aga cgg Leu Leu Arg Arg	g ctg gag 768 g Leu Glu 255					
gac gag gag ga Asp Glu Glu As 26	sp Asp Asp Al	g acg cag aaa a Thr Gln Lys 265	agc ggc ggc tac Ser Gly Gly Ty: 27	Val Asp					
cgg ttt ctc to Arg Phe Leu Se 275				843					
<210> 2 <211> 280 <212> PRT <213> Hansenula polymorpha									
<400> 2									
Met Asp Phe Me	et Lys Pro Gl 5	u Thr Val Leu 10	Asp Leu Gly As:	n Ile Arg 15					
Asp Ala Leu Va		u Asp Thr Ile 25	: Ile Phe Asn Ph 30	e Ile Glu					

Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe 35

Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln 55 His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu 65 70 75 80 Val Pro Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn 105 Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala 120 Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala 135 Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg 150 155 Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp 170 165 Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn 185 Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg 200 Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser 215 Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile 230 235 Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu 250 Asp Glu Glu Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp 265 Arg Phe Leu Ser Ser Gly Leu Tyr <210> 3 <211> 1655

<212> DNA

<213> Hansenula polymorpha

<220>

<221> gene

(1)..(1655) <222>

<223> 1,8 kb genomic DNA-fragment from Hanseula polymorpha

<220>

<221> gene

(1)..(1655) <222>

<223> 1,8 kb genomic DNA-fragment from Hansenula polymorpha

<400> 3 cccggcccaa tgccagcaat atggagacgt ttaggcagaa taggcgttcc atacttctca 60 cgctgcttgt tgccaccgga atatacaccg cattgcagtt tgcacacatc atactatatg 120 acgattacat tggcggaacg tatcgcgagt cgctcacgag acgcattaga atgacagaga 180 aatcgcgaaa cgaccttata gacgcacgtg aaaactacgg gtttggaggc agcaaggagg 240 agcgaatcca gcggtttttg tggttcagac atctttcgtg gcttttaggc gaggataagc 300 gaacttgagg agcgtttttt ttttcctgtt tagtttttgt aggtatggac tttatgaagc 360 cagaaacagt gctggacctt ggcaacatta gagatgcctt ggtccggatg gaggatacga 420 tcatcttcaa ctttatcgag cggtcgcagt tctatgcgtc gccctcggta tacaaagtca 480 accagtteee tatteeeaac ttegaegget egttettgga etggetgttg tegeageaeg 540 agcgaatcca ttcgcaggtg aggagatacg acgcgccaga cgaggtgcct tttttcccca 600 acgtgctgga aaaaacgttt ctgcccaaga tcaactaccc atcggtgcta gcctcctacg 660 cggatgaaat caacgtcaac aaagagatac tcaagatcta cacgtcagag atagtaccag 720 gaatagctgc aggcagcgga gagcaggagg acaaccttgg ctcgtgcgca atggccqaca 780 togagtgcct gcagtcgcta tocagaagaa tocattttgg ccgttttgtc gcagaggcta 840 aatttatcag tgagggggac aagattgtgg atctgatcaa aaagagagat gtggaaggca 900 ttgaggcgct catcacaaac gccgaggtcg aaaaacggat cttggacaga cttctggaga 960 agggaagggc gtatggaaca gacccgacac taaagttcac gcagcacatt cagagcaagg 1020 tgaagcccga ggtgattgtg aaaatctaca aggatttcgt gattccgctc acgaagaagg 1080 tcgaagtcga ctacttgctg agacggctgg aggacgagga ggacgatgat gcgacgcaga 1140 aaagcggcgg ctacgttgac cggtttctct cctctggctt gtactagaaa ttaaaatttt 1200 cagtacttta attattctcg aattctagtt cagataccgc atggtaattt caaaggccag 1260 aaaagtggcc gcgttggctg gggcagctct cagaatagtc ggcgagaatc ctttgactag 1320 eccccaggea eegetetgte tecaaataee eetaatagte teaacageat ttetataaae 1380 cagcttcttg tagttgtccg tctgcatgtt ggacttgatc acatcgatcg gataaatact 1440

gaaccacat	ccgtaacctg	ccagcgcccc	aaagacgcag	agcttccagt	tctcgatgtc	1500
cttcctggca	a atattccgcg	actcgatctc	gtttttcacg	agagcttcaa	aagtcagaaa	1560
atacgctcc	g ctacccaaac	tttctcttgc	cagcgtaggt	cccagacccc	ggtagattaa	1620
cttgatgcct	cccgtatggt	acagcttctt	gatcc			1655
<210> 4 <211> 20 <212> DNZ <213> Art	<u>a</u> cificial					
<220> <223> Oli	igonucleotide	2				
<400> 4 aattaaccct	cactaaaggg					20
<210> 5 <211> 22 <212> DNZ <213> Art	<u>\</u> ificial					
<220> <223> Oli	gonucleotide	<u>2</u>				
<400> 5 gtaatacgac	: tcactatagg	gc				22
<210> 6 <211> 26 <212> DNA <213> Art	ificial					
<220> <223> Oli	gonucleotide	2				
<400> 6 atatagatet	acaaaaacta	aacagg				26
<210> 7 <211> 28 <212> DNA <213> Art	ificial					
<220> <223> Oli	gonucleotide	!				
<400> 7	gatgcgacgc	agaaaagc				28
						